



# 5

## SEQUENCE LISTING

Millennium Pharmaceuticals, Inc.  
Rosana Kapeller-Libermann

<120> 23430, A NOVEL HUMAN UBIQUITIN HYDROLASE  
FAMILY MEMBER AND USES THEREFOR

<130> 38155-20024.00

<140> US 09/905,301

<141> 2001-07-13

<150> US 60/218,245

<151> 2000-07-14

<160> 8

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 4428

<212> DNA

<213> Homo sapiens

<220>

<221> CDS

<222> (589)...(3717)

<400> 1

```
gcgcccgcgg tccttcccgg aggttccttag gtccttcact tcctgtttgc cggcacttca      60
aaatccgggt caaagggtgt cgcttcgggc tcttctcccc ggctgatccc agcactctcc      120
gtgacagcgc ctctgactc agcccaggac cggcttcttc tcacgacctg ctggagactg      180
gacgcccaca cctgaccggg aactcggagg cgtgcttctt ccacccgcgg gctagcagcc      240
ccggggccctg agctcccggc gacgcccgtg gggggcccga caggcccctc ggcgctgatg      300
ctgagtggga tcgagggccc ggggcccggg cggagtagcg gcctctggcg ccttaggcca      360
gccgcaggtg tcggttctta ggctctccag gctcgttagc tccgcccccg gcttggatgg      420
gtctccctgc gccataaatg tggctgctga ggcggcgggt gccgtggccc gtcgcgctgc      480
tgctgcggcg ctccaagttc atctccgccc cggggctctc ctgccccacc tcggggctgc      540
cgccaccgcg tccttatccc ctggccctgg ccttgacagc tggcgaca atg gac aag      597
                                     Met Asp Lys
```

1

```
atc ctg gag ggc ctt gtg agt tcc tcg cat ccc ctg ccc ctc aag cgg      645
Ile Leu Glu Gly Leu Val Ser Ser Ser His Pro Leu Pro Leu Lys Arg
      5              10              15
```

```
gtg att gtg cgg aag gtg gtg gaa tcg gcg gag cac tgg cta gac gag      693
Val Ile Val Arg Lys Val Val Glu Ser Ala Glu His Trp Leu Asp Glu
      20              25              30              35
```

```
gcg cag tgc gag gcc atg ttt gac ctg acg acc cgg ctc atc ctg gag      741
Ala Gln Cys Glu Ala Met Phe Asp Leu Thr Thr Arg Leu Ile Leu Glu
      40              45              50
```

```
ggc cag gac cct ttc cag cgg cag gtg ggg cac cag gtg ctg gag gcc      789
```

Gly	Gln	Asp	Pro	Phe	Gln	Arg	Gln	Val	Gly	His	Gln	Val	Leu	Glu	Ala		
			55					60					65				
tac	gca	cga	tac	cac	cgg	cca	gag	ttc	gag	tcc	ttc	ttc	aac	aag	acc	837	
Tyr	Ala	Arg	Tyr	His	Arg	Pro	Glu	Phe	Glu	Ser	Phe	Phe	Asn	Lys	Thr		
		70					75					80					
ttc	gtg	ttg	ggc	ctc	ctt	cat	cag	ggc	tac	cac	tct	ctg	gac	agg	aag	885	
Phe	Val	Leu	Gly	Leu	Leu	His	Gln	Gly	Tyr	His	Ser	Leu	Asp	Arg	Lys		
	85					90					95						
gat	gta	gcc	atc	ctg	gac	tac	att	cac	aac	ggc	ctg	aag	ctg	att	atg	933	
Asp	Val	Ala	Ile	Leu	Asp	Tyr	Ile	His	Asn	Gly	Leu	Lys	Leu	Ile	Met		
100					105					110					115		
agc	tgt	ccg	tcg	gtg	ctg	gat	ctc	ttt	agc	ctc	ctg	cag	gta	gag	gtg	981	
Ser	Cys	Pro	Ser	Val	Leu	Asp	Leu	Phe	Ser	Leu	Leu	Gln	Val	Glu	Val		
				120					125					130			
tta	cgg	atg	gtg	tgt	gag	agg	ccg	gag	ccg	cag	ctc	tgt	gcc	cga	ctg	1029	
Leu	Arg	Met	Val	Cys	Glu	Arg	Pro	Glu	Pro	Gln	Leu	Cys	Ala	Arg	Leu		
			135					140					145				
agc	gac	ctt	ctg	acc	gac	ttt	gtg	caa	tgc	atc	ccc	aag	ggg	aaa	ttg	1077	
Ser	Asp	Leu	Leu	Thr	Asp	Phe	Val	Gln	Cys	Ile	Pro	Lys	Gly	Lys	Leu		
		150					155					160					
tcc	atc	acg	ttc	tgt	caa	cag	ctg	gtt	cga	acg	ata	ggc	cat	ttc	cag	1125	
Ser	Ile	Thr	Phe	Cys	Gln	Gln	Leu	Val	Arg	Thr	Ile	Gly	His	Phe	Gln		
	165					170					175						
tgc	gtg	tcc	acc	cag	gaa	aga	gag	ctg	cgg	gaa	tat	gtc	tcc	cag	gtg	1173	
Cys	Val	Ser	Thr	Gln	Glu	Arg	Glu	Leu	Arg	Glu	Tyr	Val	Ser	Gln	Val		
180					185					190					195		
aca	aaa	gtg	agt	aac	ttg	ctg	cag	aac	atc	tgg	aag	gcc	gag	cct	gcc	1221	
Thr	Lys	Val	Ser	Asn	Leu	Leu	Gln	Asn	Ile	Trp	Lys	Ala	Glu	Pro	Ala		
				200					205				210				
aca	cta	ctg	cct	tcc	ctg	caa	gaa	gtt	ttt	gca	agc	atc	tct	tcc	aca	1269	
Thr	Leu	Leu	Pro	Ser	Leu	Gln	Glu	Val	Phe	Ala	Ser	Ile	Ser	Ser	Thr		
			215					220					225				
gat	gca	tca	ttt	gaa	cct	tct	gta	gca	ttg	gca	agc	ctt	gtg	cag	cat	1317	
Asp	Ala	Ser	Phe	Glu	Pro	Ser	Val	Ala	Leu	Ala	Ser	Leu	Val	Gln	His		
		230					235					240					
att	cct	ctt	cag	atg	att	aca	gtt	ctc	atc	agg	agc	ctt	act	acg	gat	1365	
Ile	Pro	Leu	Gln	Met	Ile	Thr	Val	Leu	Ile	Arg	Ser	Leu	Thr	Thr	Asp		
	245					250					255						
cca	aat	gta	aaa	gat	gca	agt	atg	acc	caa	gcc	ctt	tgc	aga	atg	att	1413	
Pro	Asn	Val	Lys	Asp	Ala	Ser	Met	Thr	Gln	Ala	Leu	Cys	Arg	Met	Ile		
260					265					270					275		
gac	tgg	cta	tcc	tgg	cca	ttg	gct	cag	cat	gtg	gat	aca	tgg	gta	att	1461	
Asp	Trp	Leu	Ser	Trp	Pro	Leu	Ala	Gln	His	Val	Asp	Thr	Trp	Val	Ile		

280										285					290					
gca	ctc	ctg	aaa	gga	ctg	gca	gct	gtc	cag	aag	ttt	act	att	ttg	ata	1509				
Ala	Leu	Leu	Lys	Gly	Leu	Ala	Ala	Val	Gln	Lys	Phe	Thr	Ile	Leu	Ile					
			295				300				305									
gat	gtt	act	ttg	ctg	aaa	ata	gaa	ctg	gtt	ttt	aat	cga	ctt	tggt	ttt	1557				
Asp	Val	Thr	Leu	Leu	Lys	Ile	Glu	Leu	Val	Phe	Asn	Arg	Leu	Trp	Phe					
			310				315				320									
cct	ctt	gtg	aga	cct	ggt	gct	ctt	gca	gtt	ctt	tct	cac	atg	ctg	ctt	1605				
Pro	Leu	Val	Arg	Pro	Gly	Ala	Leu	Ala	Val	Leu	Ser	His	Met	Leu	Leu					
			325				330				335									
agc	ttt	cag	cat	tct	cca	gag	gcg	ttc	cat	ttg	att	gtt	cct	cat	gtg	1653				
Ser	Phe	Gln	His	Ser	Pro	Glu	Ala	Phe	His	Leu	Ile	Val	Pro	His	Val					
			340				345				350				355					
gtt	aat	ttg	gtt	cat	tct	ttc	aaa	aat	gat	ggt	ctg	cct	tca	agt	aca	1701				
Val	Asn	Leu	Val	His	Ser	Phe	Lys	Asn	Asp	Gly	Leu	Pro	Ser	Ser	Thr					
			360				365				370									
gcc	ttc	tta	gta	caa	tta	aca	gaa	ttg	ata	cac	tgt	atg	atg	tat	cat	1749				
Ala	Phe	Leu	Val	Gln	Leu	Thr	Glu	Leu	Ile	His	Cys	Met	Met	Tyr	His					
			375				380				385									
tat	tct	gga	ttt	cca	gat	ctc	tat	gaa	cct	att	ctg	gag	gca	ata	aag	1797				
Tyr	Ser	Gly	Phe	Pro	Asp	Leu	Tyr	Glu	Pro	Ile	Leu	Glu	Ala	Ile	Lys					
			390				395				400									
gat	ttt	cct	aag	ccc	agt	gaa	gag	aag	att	aag	tta	att	ctc	aat	caa	1845				
Asp	Phe	Pro	Lys	Pro	Ser	Glu	Glu	Lys	Ile	Lys	Leu	Ile	Leu	Asn	Gln					
			405				410				415									
agt	gcc	tggt	act	tct	caa	tcc	aat	tct	ttg	gcg	tct	tgc	ttg	tct	aga	1893				
Ser	Ala	Trp	Thr	Ser	Gln	Ser	Asn	Ser	Leu	Ala	Ser	Cys	Leu	Ser	Arg					
			420				425				430				435					
ctt	tct	gga	aaa	tct	gaa	act	ggg	aaa	act	ggt	ctt	att	aac	cta	gga	1941				
Leu	Ser	Gly	Lys	Ser	Glu	Thr	Gly	Lys	Thr	Gly	Leu	Ile	Asn	Leu	Gly					
			440				445				450									
aat	aca	tgt	tat	atg	aac	agt	gtt	ata	caa	gcc	ttg	ttt	atg	gcc	aca	1989				
Asn	Thr	Cys	Tyr	Met	Asn	Ser	Val	Ile	Gln	Ala	Leu	Phe	Met	Ala	Thr					
			455				460				465									
gat	ttc	agg	aga	caa	gta	tta	tct	tta	aat	cta	aat	ggg	tgc	aat	tca	2037				
Asp	Phe	Arg	Arg	Gln	Val	Leu	Ser	Leu	Asn	Leu	Asn	Gly	Cys	Asn	Ser					
			470				475				480									
tta	atg	aaa	aaa	tta	cag	cat	ctt	ttt	gcc	ttt	ctg	gcc	cat	aca	cag	2085				
Leu	Met	Lys	Lys	Leu	Gln	His	Leu	Phe	Ala	Phe	Leu	Ala	His	Thr	Gln					
			485				490				495									
agg	gaa	gca	tac	gca	cct	cgg	ata	ttc	ttt	gag	gct	tcc	aga	cct	cca	2133				
Arg	Glu	Ala	Tyr	Ala	Pro	Arg	Ile	Phe	Phe	Glu	Ala	Ser	Arg	Pro	Pro					
			500				505				510				515					

tgg ttt act ccc aga tca cag caa gac tgt tct gaa tac ctc aga ttt	2181
Trp Phe Thr Pro Arg Ser Gln Gln Asp Cys Ser Glu Tyr Leu Arg Phe	
520 525 530	
ctc ctt gac agg ctc cat gaa gaa gaa aag atc ttg aaa gtt cag gcc	2229
Leu Leu Asp Arg Leu His Glu Glu Glu Lys Ile Leu Lys Val Gln Ala	
535 540 545	
tca cac aag cct tct gaa att ctg gaa tgc agt gaa act tct tta cag	2277
Ser His Lys Pro Ser Glu Ile Leu Glu Cys Ser Glu Thr Ser Leu Gln	
550 555 560	
gaa gta gct agt aaa gca gca gta cta aca gag acc cct cgt aca agt	2325
Glu Val Ala Ser Lys Ala Ala Val Leu Thr Glu Thr Pro Arg Thr Ser	
565 570 575	
gac ggt gag aag act tta ata gaa aaa atg ttt gga gga aaa cta cga	2373
Asp Gly Glu Lys Thr Leu Ile Glu Lys Met Phe Gly Gly Lys Leu Arg	
580 585 590 595	
act cac ata cgt tgt ttg aac tgc agg agt acc tca caa aaa gtg gaa	2421
Thr His Ile Arg Cys Leu Asn Cys Arg Ser Thr Ser Gln Lys Val Glu	
600 605 610	
gcc ttt aca gat ctt tcg ctt gcc ttt tgt cct tcc tct tct ttg gaa	2469
Ala Phe Thr Asp Leu Ser Leu Ala Phe Cys Pro Ser Ser Ser Leu Glu	
615 620 625	
aac atg tct gtc caa gat cca gca tca tca ccc agt ata caa gat ggt	2517
Asn Met Ser Val Gln Asp Pro Ala Ser Ser Pro Ser Ile Gln Asp Gly	
630 635 640	
ggt cta atg caa gcc tct gta ccc ggt cct tca gaa gaa cca gta gtt	2565
Gly Leu Met Gln Ala Ser Val Pro Gly Pro Ser Glu Glu Pro Val Val	
645 650 655	
tat aat cca aca aca gct gcc ttc atc tgt gac tca ctt gtg aat gaa	2613
Tyr Asn Pro Thr Thr Ala Ala Phe Ile Cys Asp Ser Leu Val Asn Glu	
660 665 670 675	
aaa acc ata ggc agt cct cct aat gag ttt tac tgt tct gaa aac act	2661
Lys Thr Ile Gly Ser Pro Pro Asn Glu Phe Tyr Cys Ser Glu Asn Thr	
680 685 690	
tct gtc cct aac gaa tct aac aag att ctt gtt aat aaa gat gta cct	2709
Ser Val Pro Asn Glu Ser Asn Lys Ile Leu Val Asn Lys Asp Val Pro	
695 700 705	
cag aaa cca gga ggt gaa acc aca cct tca gta act gac tta cta aat	2757
Gln Lys Pro Gly Gly Glu Thr Thr Pro Ser Val Thr Asp Leu Leu Asn	
710 715 720	
tat ttt ttg gct cca gag att ctt act ggt gat aac caa tat tat tgt	2805
Tyr Phe Leu Ala Pro Glu Ile Leu Thr Gly Asp Asn Gln Tyr Tyr Cys	
725 730 735	

gaa aac tgt gcc tct ctg caa aat gct gag aaa act atg caa atc acg Glu Asn Cys Ala Ser Leu Gln Asn Ala Glu Lys Thr Met Gln Ile Thr 740 745 750 755	2853
gag gaa cct gaa tac ctt att ctt act ctc ctg aga ttt tca tat gat Glu Glu Pro Glu Tyr Leu Ile Leu Thr Leu Leu Arg Phe Ser Tyr Asp 760 765 770	2901
cag aag tat cat gtg aga agg aaa att tta gac aat gta tca ctg cca Gln Lys Tyr His Val Arg Arg Lys Ile Leu Asp Asn Val Ser Leu Pro 775 780 785	2949
ctg gtt ttg gag ttg cca gtt aaa aga att act tct ttc tct tca ttg Leu Val Leu Glu Leu Pro Val Lys Arg Ile Thr Ser Phe Ser Ser Leu 790 795 800	2997
tca gaa agt tgg tct gta gat gtt gac ttc act gat ctt agt gag aac Ser Glu Ser Trp Ser Val Asp Val Asp Phe Thr Asp Leu Ser Glu Asn 805 810 815	3045
ctt gct aaa aaa tta aag cct tca ggg act gat gaa gct tcc tgc aca Leu Ala Lys Lys Leu Lys Pro Ser Gly Thr Asp Glu Ala Ser Cys Thr 820 825 830 835	3093
aaa ttg gtg ccc tat cta tta agt tcc gtt gtg gtt cac tct ggt ata Lys Leu Val Pro Tyr Leu Leu Ser Ser Val Val Val His Ser Gly Ile 840 845 850	3141
tcc tct gaa agt ggg cat tac tat tct tat gcc aga aat atc aca agt Ser Ser Glu Ser Gly His Tyr Tyr Ser Tyr Ala Arg Asn Ile Thr Ser 855 860 865	3189
aca gac tct tca tat cag atg tac cac cag tct gag gct ctg gca tta Thr Asp Ser Ser Tyr Gln Met Tyr His Gln Ser Glu Ala Leu Ala Leu 870 875 880	3237
gca tcc tcc cag agt cat tta cta ggg aga gat agt ccc agt gca gtt Ala Ser Ser Gln Ser His Leu Leu Gly Arg Asp Ser Pro Ser Ala Val 885 890 895	3285
ttt gaa cag gat ttg gaa aat aag gaa atg tca aaa gaa tgg ttt tta Phe Glu Gln Asp Leu Glu Asn Lys Glu Met Ser Lys Glu Trp Phe Leu 900 905 910 915	3333
ttt aat gac agt aga gtg aca ttt act tca ttt cag tca gtc cag aaa Phe Asn Asp Ser Arg Val Thr Phe Thr Ser Phe Gln Ser Val Gln Lys 920 925 930	3381
att acg agc agg ttt cca aag gac aca gct tat gtg ctt ttg tat aaa Ile Thr Ser Arg Phe Pro Lys Asp Thr Ala Tyr Val Leu Leu Tyr Lys 935 940 945	3429
aaa cag cat agt act aat ggt tta agt ggt aat aac cca acc agt gga Lys Gln His Ser Thr Asn Gly Leu Ser Gly Asn Asn Pro Thr Ser Gly 950 955 960	3477
ctc tgg ata aat gga gac cca cct cta cag aaa gaa ctt atg gat gct	3525

Leu Trp Ile Asn Gly Asp Pro Pro Leu Gln Lys Glu Leu Met Asp Ala  
 965 970 975

ata aca aaa gac aat aaa cta tat tta cag gaa caa gag ttg aat gct 3573  
 Ile Thr Lys Asp Asn Lys Leu Tyr Leu Gln Glu Gln Glu Leu Asn Ala  
 980 985 990 995

cga gcc cgg gcc ctc caa gct gca tct gct tca tgt tca ttt cgg ccc 3621  
 Arg Ala Arg Ala Leu Gln Ala Ala Ser Ala Ser Cys Ser Phe Arg Pro  
 1000 1005 1010

aat gga ttt gat gac aac gac cca cca gga agc tgt gga cca act ggt 3669  
 Asn Gly Phe Asp Asp Asn Asp Pro Pro Gly Ser Cys Gly Pro Thr Gly  
 1015 1020 1025

gga ggg ggt gga gga gga ttt aat aca gtt ggc aga ctc gta ttt tga 3717  
 Gly Gly Gly Gly Gly Gly Phe Asn Thr Val Gly Arg Leu Val Phe \*  
 1030 1035 1040

tcctgagaga gtccaaaatg cactgggtcac gaaacgtcta atactatgac tgttaaaatg 3777  
 tcagactata acaaataatct atctttttatt ttctcattaga cccttataact tcaagagAAC 3837  
 acactcagtg cttgtttttta ttttcttgac acattttatta acaaaatgca tcatggaaaa 3897  
 aaaatctacc tcttaaaatt ccatttgctt ttatgggttag acatgcttga ccaaaaatgt 3957  
 tcagaagaaa atatgtacct ggtccctaata taagctgcgt taaatttggt agaagcattt 4017  
 aaatggtcta tcttcagttt tactgaacaa aaaatgtaata ttatttagca ttctttataa 4077  
 aagaattgat gctagaggta aaaaaaataa cttgtttttta aaaaatcctt tacgtcttgt 4137  
 gtaattaccc cattattaaa ttcaagtcct tgaaaatcaa ctagagatta taaagtcctt 4197  
 aaagaaggca ataacaaaat ttatcaagat atagtacttt tcagtttttg tttagtgtct 4257  
 tcagcatcac tgtgtctgta tttcaagtac aaatgttttt aaaaaggatt ctttatacat 4317  
 atgtgctgaa ttgattttta ggaaagttgc atgatcctgt aggagcaaca tttttaccta 4377  
 aaaaatgcta actttatagt atttctaatt gttcaaggat tttaaaattc t 4428

<210> 2  
 <211> 1042  
 <212> PRT  
 <213> Homo sapiens

<400> 2

Met Asp Lys Ile Leu Glu Gly Leu Val Ser Ser Ser His Pro Leu Pro  
 1 5 10 15  
 Leu Lys Arg Val Ile Val Arg Lys Val Val Glu Ser Ala Glu His Trp  
 20 25 30  
 Leu Asp Glu Ala Gln Cys Glu Ala Met Phe Asp Leu Thr Thr Arg Leu  
 35 40 45  
 Ile Leu Glu Gly Gln Asp Pro Phe Gln Arg Gln Val Gly His Gln Val  
 50 55 60  
 Leu Glu Ala Tyr Ala Arg Tyr His Arg Pro Glu Phe Glu Ser Phe Phe  
 65 70 75 80  
 Asn Lys Thr Phe Val Leu Gly Leu Leu His Gln Gly Tyr His Ser Leu  
 85 90 95  
 Asp Arg Lys Asp Val Ala Ile Leu Asp Tyr Ile His Asn Gly Leu Lys  
 100 105 110  
 Leu Ile Met Ser Cys Pro Ser Val Leu Asp Leu Phe Ser Leu Leu Gln  
 115 120 125  
 Val Glu Val Leu Arg Met Val Cys Glu Arg Pro Glu Pro Gln Leu Cys  
 130 135 140  
 Ala Arg Leu Ser Asp Leu Leu Thr Asp Phe Val Gln Cys Ile Pro Lys

145					150					155					160
Gly	Lys	Leu	Ser	Ile	Thr	Phe	Cys	Gln	Gln	Leu	Val	Arg	Thr	Ile	Gly
				165					170					175	
His	Phe	Gln	Cys	Val	Ser	Thr	Gln	Glu	Arg	Glu	Leu	Arg	Glu	Tyr	Val
			180					185					190		
Ser	Gln	Val	Thr	Lys	Val	Ser	Asn	Leu	Leu	Gln	Asn	Ile	Trp	Lys	Ala
		195					200				205				
Glu	Pro	Ala	Thr	Leu	Leu	Pro	Ser	Leu	Gln	Glu	Val	Phe	Ala	Ser	Ile
	210					215				220					
Ser	Ser	Thr	Asp	Ala	Ser	Phe	Glu	Pro	Ser	Val	Ala	Leu	Ala	Ser	Leu
225				230				235						240	
Val	Gln	His	Ile	Pro	Leu	Gln	Met	Ile	Thr	Val	Leu	Ile	Arg	Ser	Leu
			245					250					255		
Thr	Thr	Asp	Pro	Asn	Val	Lys	Asp	Ala	Ser	Met	Thr	Gln	Ala	Leu	Cys
		260					265					270			
Arg	Met	Ile	Asp	Trp	Leu	Ser	Trp	Pro	Leu	Ala	Gln	His	Val	Asp	Thr
	275						280				285				
Trp	Val	Ile	Ala	Leu	Leu	Lys	Gly	Leu	Ala	Ala	Val	Gln	Lys	Phe	Thr
	290					295				300					
Ile	Leu	Ile	Asp	Val	Thr	Leu	Leu	Lys	Ile	Glu	Leu	Val	Phe	Asn	Arg
305				310				315						320	
Leu	Trp	Phe	Pro	Leu	Val	Arg	Pro	Gly	Ala	Leu	Ala	Val	Leu	Ser	His
			325					330					335		
Met	Leu	Leu	Ser	Phe	Gln	His	Ser	Pro	Glu	Ala	Phe	His	Leu	Ile	Val
		340					345					350			
Pro	His	Val	Val	Asn	Leu	Val	His	Ser	Phe	Lys	Asn	Asp	Gly	Leu	Pro
	355					360					365				
Ser	Ser	Thr	Ala	Phe	Leu	Val	Gln	Leu	Thr	Glu	Leu	Ile	His	Cys	Met
	370				375					380					
Met	Tyr	His	Tyr	Ser	Gly	Phe	Pro	Asp	Leu	Tyr	Glu	Pro	Ile	Leu	Glu
385				390				395						400	
Ala	Ile	Lys	Asp	Phe	Pro	Lys	Pro	Ser	Glu	Glu	Lys	Ile	Lys	Leu	Ile
			405					410					415		
Leu	Asn	Gln	Ser	Ala	Trp	Thr	Ser	Gln	Ser	Asn	Ser	Leu	Ala	Ser	Cys
		420					425					430			
Leu	Ser	Arg	Leu	Ser	Gly	Lys	Ser	Glu	Thr	Gly	Lys	Thr	Gly	Leu	Ile
	435					440				445					
Asn	Leu	Gly	Asn	Thr	Cys	Tyr	Met	Asn	Ser	Val	Ile	Gln	Ala	Leu	Phe
	450			455				460							
Met	Ala	Thr	Asp	Phe	Arg	Arg	Gln	Val	Leu	Ser	Leu	Asn	Leu	Asn	Gly
465				470				475					480		
Cys	Asn	Ser	Leu	Met	Lys	Lys	Leu	Gln	His	Leu	Phe	Ala	Phe	Leu	Ala
			485					490					495		
His	Thr	Gln	Arg	Glu	Ala	Tyr	Ala	Pro	Arg	Ile	Phe	Phe	Glu	Ala	Ser
		500					505					510			
Arg	Pro	Pro	Trp	Phe	Thr	Pro	Arg	Ser	Gln	Gln	Asp	Cys	Ser	Glu	Tyr
	515					520					525				
Leu	Arg	Phe	Leu	Leu	Asp	Arg	Leu	His	Glu	Glu	Glu	Lys	Ile	Leu	Lys
	530				535					540					
Val	Gln	Ala	Ser	His	Lys	Pro	Ser	Glu	Ile	Leu	Glu	Cys	Ser	Glu	Thr
545				550				555						560	
Ser	Leu	Gln	Glu	Val	Ala	Ser	Lys	Ala	Ala	Val	Leu	Thr	Glu	Thr	Pro
			565					570					575		
Arg	Thr	Ser	Asp	Gly	Glu	Lys	Thr	Leu	Ile	Glu	Lys	Met	Phe	Gly	Gly
		580					585					590			
Lys	Leu	Arg	Thr	His	Ile	Arg	Cys	Leu	Asn	Cys	Arg	Ser	Thr	Ser	Gln
	595					600					605				

Lys	Val	Glu	Ala	Phe	Thr	Asp	Leu	Ser	Leu	Ala	Phe	Cys	Pro	Ser	Ser	
610						615					620					
Ser	Leu	Glu	Asn	Met	Ser	Val	Gln	Asp	Pro	Ala	Ser	Ser	Pro	Ser	Ile	
625					630					635					640	
Gln	Asp	Gly	Gly	Leu	Met	Gln	Ala	Ser	Val	Pro	Gly	Pro	Ser	Glu	Glu	
				645					650					655		
Pro	Val	Val	Tyr	Asn	Pro	Thr	Thr	Ala	Ala	Phe	Ile	Cys	Asp	Ser	Leu	
			660					665					670			
Val	Asn	Glu	Lys	Thr	Ile	Gly	Ser	Pro	Pro	Asn	Glu	Phe	Tyr	Cys	Ser	
		675				680						685				
Glu	Asn	Thr	Ser	Val	Pro	Asn	Glu	Ser	Asn	Lys	Ile	Leu	Val	Asn	Lys	
	690				695					700						
Asp	Val	Pro	Gln	Lys	Pro	Gly	Gly	Glu	Thr	Thr	Pro	Ser	Val	Thr	Asp	
705				710						715					720	
Leu	Leu	Asn	Tyr	Phe	Leu	Ala	Pro	Glu	Ile	Leu	Thr	Gly	Asp	Asn	Gln	
				725					730					735		
Tyr	Tyr	Cys	Glu	Asn	Cys	Ala	Ser	Leu	Gln	Asn	Ala	Glu	Lys	Thr	Met	
			740					745					750			
Gln	Ile	Thr	Glu	Glu	Pro	Glu	Tyr	Leu	Ile	Leu	Thr	Leu	Leu	Arg	Phe	
		755					760					765				
Ser	Tyr	Asp	Gln	Lys	Tyr	His	Val	Arg	Arg	Lys	Ile	Leu	Asp	Asn	Val	
	770					775					780					
Ser	Leu	Pro	Leu	Val	Leu	Glu	Leu	Pro	Val	Lys	Arg	Ile	Thr	Ser	Phe	
785					790					795					800	
Ser	Ser	Leu	Ser	Glu	Ser	Trp	Ser	Val	Asp	Val	Asp	Phe	Thr	Asp	Leu	
				805					810					815		
Ser	Glu	Asn	Leu	Ala	Lys	Lys	Leu	Lys	Pro	Ser	Gly	Thr	Asp	Glu	Ala	
		820						825					830			
Ser	Cys	Thr	Lys	Leu	Val	Pro	Tyr	Leu	Leu	Ser	Ser	Val	Val	Val	His	
		835					840					845				
Ser	Gly	Ile	Ser	Ser	Glu	Ser	Gly	His	Tyr	Tyr	Ser	Tyr	Ala	Arg	Asn	
	850					855					860					
Ile	Thr	Ser	Thr	Asp	Ser	Ser	Tyr	Gln	Met	Tyr	His	Gln	Ser	Glu	Ala	
865					870					875					880	
Leu	Ala	Leu	Ala	Ser	Ser	Gln	Ser	His	Leu	Leu	Gly	Arg	Asp	Ser	Pro	
				885					890					895		
Ser	Ala	Val	Phe	Glu	Gln	Asp	Leu	Glu	Asn	Lys	Glu	Met	Ser	Lys	Glu	
			900					905					910			
Trp	Phe	Leu	Phe	Asn	Asp	Ser	Arg	Val	Thr	Phe	Thr	Ser	Phe	Gln	Ser	
	915						920						925			
Val	Gln	Lys	Ile	Thr	Ser	Arg	Phe	Pro	Lys	Asp	Thr	Ala	Tyr	Val	Leu	
	930					935					940					
Leu	Tyr	Lys	Lys	Gln	His	Ser	Thr	Asn	Gly	Leu	Ser	Gly	Asn	Asn	Pro	
945					950					955					960	
Thr	Ser	Gly	Leu	Trp	Ile	Asn	Gly	Asp	Pro	Pro	Leu	Gln	Lys	Glu	Leu	
				965					970					975		
Met	Asp	Ala	Ile	Thr	Lys	Asp	Asn	Lys	Leu	Tyr	Leu	Gln	Glu	Gln	Glu	
		980						985					990			
Leu	Asn	Ala	Arg	Ala	Arg	Ala	Leu	Gln	Ala	Ala	Ser	Ala	Ser	Cys	Ser	
		995					1000					1005				
Phe	Arg	Pro	Asn	Gly	Phe	Asp	Asp	Asn	Asp	Pro	Pro	Gly	Ser	Cys	Gly	
	1010					1015					1020					
Pro	Thr	Gly	Gly	Gly	Gly	Gly	Gly	Gly	Phe	Asn	Thr	Val	Gly	Arg	Leu	
1025					1030					1035					1040	
Val	Phe															



<210> 3  
 <211> 3129  
 <212> DNA  
 <213> Homo sapiens

<400> 3  
 atggacaaga tcctggaggg ccttgtgagt tcctcgcatc ccctgcccct caagcgggtg 60  
 attgtgcgga aggtgggtgga atcggcgagg cactggctag acgaggcgca gtgcgaggcc 120  
 atgtttgacc tgacgaccgg gctcatcctg gagggccagg accctttcca gcggcagggtg 180  
 gggcaccagg tgctggaggc ctacgcacga taccaccggc cagagttcga gtcctttcttc 240  
 aacaagacct tcgtgttggg cctccttcat cagggtacc actctctgga caggaaggat 300  
 gtagccatcc tggactacat tcacaacggc ctgaagctga ttatgagctg tccgtcgggtg 360  
 ctggatctct ttagcctcct gcaggtagag gtgttacgga tgggtgtgtga gaggccggag 420  
 ccgcagctct gtgcccgaact gagcgacctt ctgaccgaact ttgtgcaatg catccccaaag 480  
 gggaaattgt ccatcacgtt ctgtcaacag ctggttcgaa cgataggcca tttccagtgc 540  
 gtgtccaccc aggaagaga gctgcgggaa tatgtctccc aggtgacaaa agtgagtaac 600  
 ttgctgcaga acatctggaa ggccgagcct gccacactac tgccttccct gcaagaagtt 660  
 tttgcaagca tctcttccac agatgcatca tttgaacctt ctgtagcatt ggcaagcctt 720  
 gtgcagcata ttcctcttca gatgattaca gttctcatca ggagccttac tacggatcca 780  
 aatgtaaaag atgcaagtat gacccaagcc ctttgcagaa tgattgactg gctatcctgg 840  
 ccattggctc agcatgtgga tacatgggta attgcactcc tgaaaggact ggcagctgtc 900  
 cagaagttaa ctattttgat agatgttact ttgctgaaaa tagaactggt ttttaatcga 960  
 ctttggtttc ctcttgtgag acctggtgct cttgcagttc tttctcacat gctgcttagc 1020  
 tttcagcatt ctccagaggc gttccatttg attgttcttc atgtggttaa tttggttcat 1080  
 tctttcaaaa atgatggtct gccttcaagt acagccttct tagtacaatt aacagaattg 1140  
 atacactgta tgatgtatca ttattctgga tttccagatc tctatgaacc tattctggag 1200  
 gcaataaagg attttcttaa gccagtgaa gagaagatta agttaattct caatcaaagt 1260  
 gcctggactt ctcaatccaa ttctttggcg tcttgcctgt ctgactttc tggaaaatct 1320  
 gaaactggga aaactggtct tattaaccta ggaaatacat gttatatgaa cagtgttata 1380  
 caagccttgt ttatggccac agatttcagg agacaagtat tatctttaa tctaaatggg 1440  
 tgcaattcat taatgaaaaa attacagcat cttttgcct ttctggccca tacacagagg 1500  
 gaagcatacg cacctcggat attctttgag gcttcagac ctccatggtt tactcccaga 1560  
 tcacagcaag actgttctga atacctcaga tttctccttg acaggctcca tgaagaagaa 1620  
 aagatcttga aagttcaggc ctcacacaag ccttctgaaa ttctggaatg cagtgaact 1680  
 tctttacagg aagtagctag taaagcagca gtactaacag agaccctcg tacaagtgac 1740  
 ggtgagaaga ctttaataga aaaaatgttt ggaggaaaac tacgaactca catacgttgt 1800  
 ttgaactgca ggagtacctc acaaaaagtg gaagccttta cagatcttct gcttgccctt 1860  
 tgccttctct cttctttgga aaacatgtct gtccaagatc cagcatcatc acccagtata 1920  
 caagatggtg gtctaattgca agcctctgta cccggtcctt cagaagaacc agtagtttat 1980  
 aatccaacaa cagctgcctt catctgtgac tcaactgtga atgaaaaaac cataggcagt 2040  
 cctcctaatt agttttactg ttctgaaaac acttctgtcc ctaacgaatc taacaagatt 2100  
 cttgttaata aagatgtacc tcagaaacca ggagggtgaa ccacaccttc agtaactgac 2160  
 ttactaaatt attttttggc tccagagatt cttactggtg ataaccaata ttattgtgaa 2220  
 aactgtgcct ctctgcaaaa tgctgagaaa actatgcaaa tcacggagga acctgaatac 2280  
 cttattctta ctctcctgag attttcatat gatcagaagt atcatgtgag aaggaaaatt 2340  
 ttagacaatg tatcactgcc actgggttttg gagttgccag ttaaaagaat tacttcttct 2400  
 tcttcattgt cagaaagttg gtctgtagat gttgacttca ctgatcttag tgagaacctt 2460  
 gctaaaaaat taaagccttc agggactgat gaagcttctt gcacaaaatt ggtgccctat 2520  
 ctattaagtt ccgttgtggt tcactctggt atatcctctg aaagtgggca ttactattct 2580  
 tatgccagaa atatcacaag tacagactct tcatatcaga tgtaccacca gtctgaggct 2640  
 ctggcattag catcctccca gagtcattta ctaggagag atagtccag tgcagttttt 2700  
 gaacaggatt tggaaaataa ggaaatgtca aaagaatggt ttttatttaa tgacagtaga 2760  
 gtgacattta cttcatttca gtcagtccag aaaattacga gcaggtttcc aaaggacaca 2820  
 gcttatgtgc ttttgtataa aaaacagcat agtactaatg gtttaagtgg taataacca 2880  
 accagtggac tctggataaa tggagacca cctctacaga aagaacttat ggatgctata 2940  
 acaaaagaca ataaactata ttacaggaa caagagtga atgctcgagc ccgggccctc 3000  
 caagctgcat ctgcttcatg ttcatttcgg ccaatggat ttgatgacaa cgaccacca 3060

ggaagctgtg gaccaactgg tggaggggggt ggaggaggat ttaatacagt tggcagactc 3120  
gtattttga 3129

<210> 4  
<211> 32  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Consensus amino acid

<400> 4  
Thr Gly Leu Ile Asn Leu Gly Asn Thr Cys Tyr Met Asn Ser Val Leu  
1 5 10 15  
Gln Cys Leu Phe Ser Ile Pro Pro Leu Arg Asp Tyr Leu Leu Asp Ile  
20 25 30

<210> 5  
<211> 69  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Consensus amino acid

<400> 5  
Gly Pro Gly Lys Tyr Glu Leu Tyr Ala Val Val Val His Ser Gly Ser  
1 5 10 15  
Ser Leu Ser Gly Gly His Tyr Thr Ala Tyr Val Lys Lys Glu Asn Trp  
20 25 30  
Tyr Lys Phe Asp Asp Asp Lys Val Ser Arg Val Thr Glu Glu Glu Val  
35 40 45  
Leu Lys Glu Ser Gly Gly Glu Ser Gly Asp Thr Ser Ser Ala Tyr Ile  
50 55 60  
Leu Phe Tyr Glu Arg  
65

<210> 6  
<211> 83  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Consensus amino acid

<400> 6  
Pro Glu Gly Asp His Ser Ser Leu Glu Asp Cys Leu Glu Gln Phe Phe  
1 5 10 15  
Lys Pro Glu Glu Leu Glu Gly Asp Asn Lys Tyr His Cys Glu Lys Cys  
20 25 30  
Lys Lys Lys Gln Asp Ala Thr Lys Gln Leu Thr Ile Lys Lys Leu Pro  
35 40 45  
Gln Val Leu Thr Ile His Leu Lys Arg Phe Glu Tyr Asn Glu Glu Arg  
50 55 60  
Phe Ser Ser Asn Lys Ile Asn Lys His Val Ser Phe Pro Leu Glu Thr  
65 70 75 80  
Leu Asp Leu

<210> 7  
 <211> 42  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Signature pattern for conserved region

<221> VARIANT  
 <222> (1)...(42)  
 <223> Xaa = Any Amino Acid

<400> 7  
 Gly Leu Ile Val Met Phe Tyr Xaa Xaa Xaa Ala Gly Cys Asn Ala Ser  
 1 5 10 15  
 Met Xaa Cys Phe Tyr Trp Leu Ile Val Met Phe Cys Asn Ser Thr Ser  
 20 25 30  
 Ala Cys Val Xaa Leu Ile Val Met Ser Gln  
 35 40

<210> 8  
 <211> 26  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Signature pattern for conserved region

<221> VARIANT  
 <222> (1)...(26)  
 <223> Xaa = Any Amino Acid

<400> 8  
 Tyr Xaa Leu Xaa Ser Ala Gly Leu Ile Val Met Phe Thr Xaa Xaa His  
 1 5 10 15  
 Xaa Gly Xaa Xaa Xaa Xaa Xaa Gly His Tyr  
 20 25